

Supplement of Hydrol. Earth Syst. Sci., 22, 3453–3472, 2018  
<https://doi.org/10.5194/hess-22-3453-2018-supplement>  
© Author(s) 2018. This work is distributed under  
the Creative Commons Attribution 3.0 License.



*Supplement of*

## **Seasonal streamflow forecasts for Europe – Part I: Hindcast verification with pseudo- and real observations**

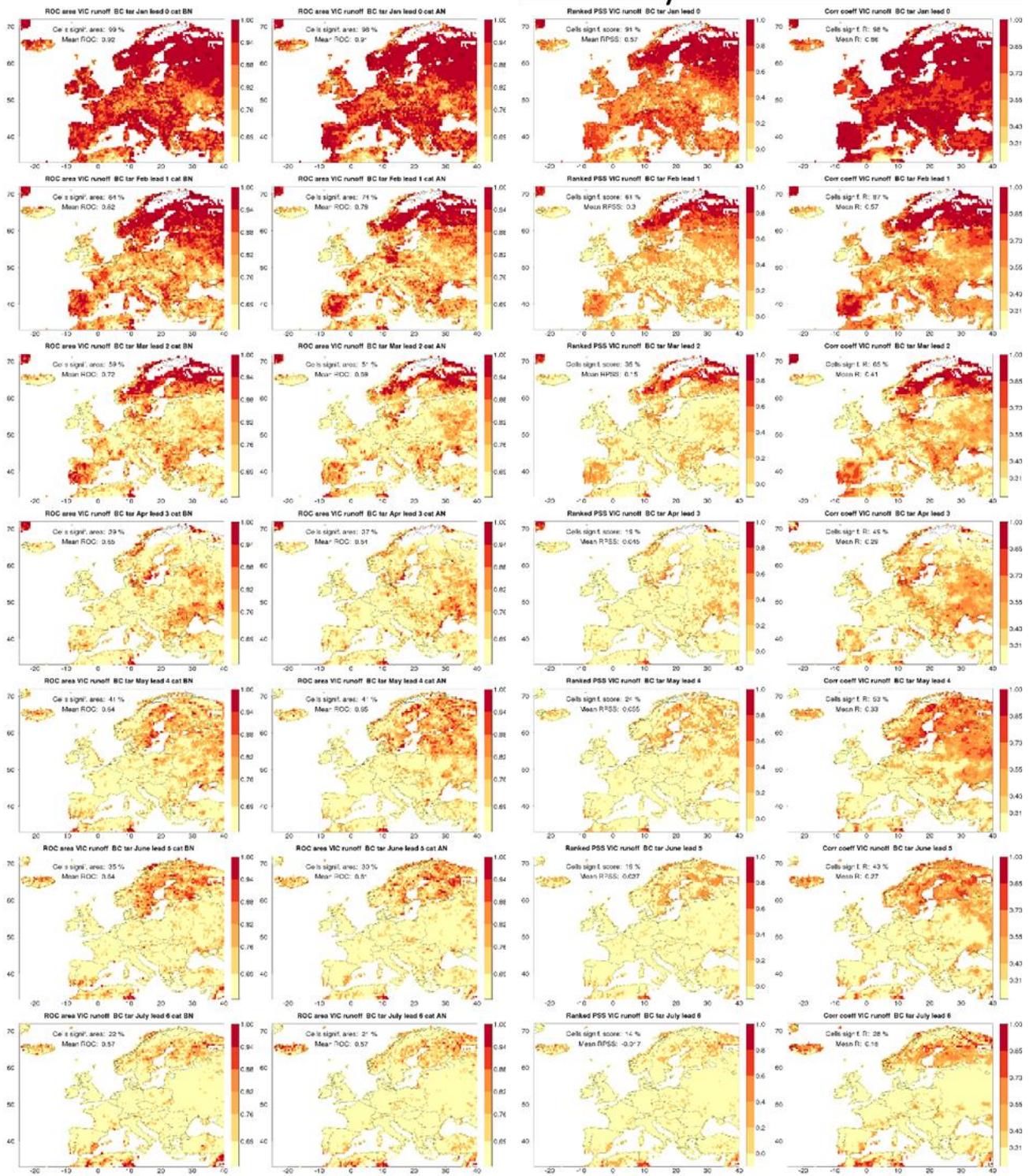
**Wouter Greuell et al.**

*Correspondence to:* Wouter Greuell ([wouter.greuell@wur.nl](mailto:wouter.greuell@wur.nl))

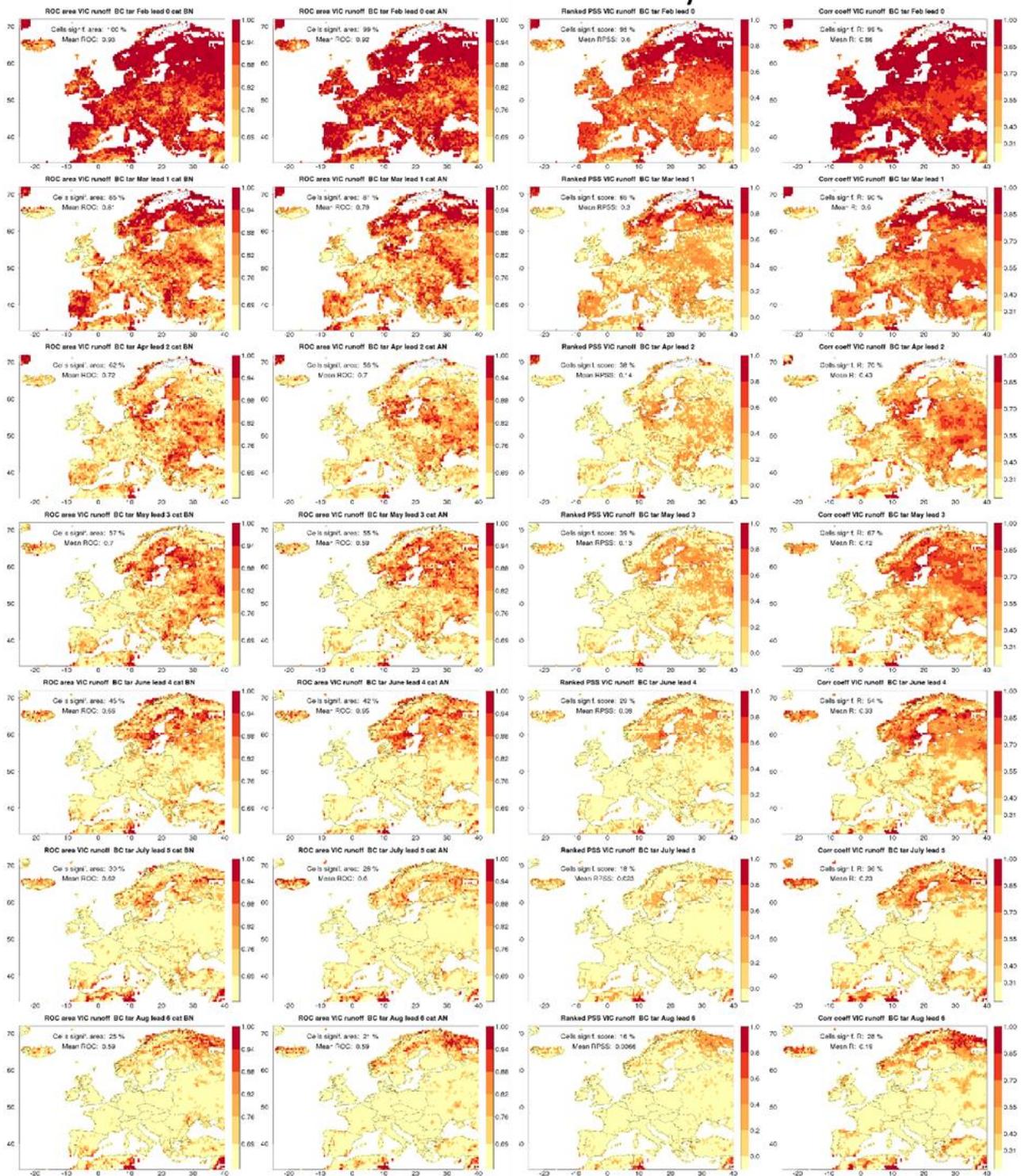
The copyright of individual parts of the supplement might differ from the CC BY 3.0 License.

Figure S1: This figure (next 12 pages) supplements Figs. 2, 3 and 9. It shows maps of skill of the runoff Full Hindcasts for all target and lead months for all four metrics. Note that each page represents all hindcasts initialised in the same month for, from top to bottom, the seven different lead months. In the right most column, skill is measured in terms of the Pearson correlation coefficient between the median of the hindcasts and the observations ( $R$ ). The threshold of significant skill lies at 0.31, so yellow cells have insignificant skill, (dark) red cells have (most) skill. White, terrestrial cells correspond to cells where observations or hindcasts consist for more than one third of zeros or one sixth of ties. The legend provides the fraction of cells with significant values of  $R$  (at the 5% level) and the domain-averaged value of  $R$ . The third column shows Ranked Probability Skill Score. Colour coding as before for  $R$ , except threshold for insignificant RPSS is 0. The second and first columns show ROC area for the high (AN) and low tercile (BN) respectively. Colour coding as before for  $R$ , except threshold for insignificant RPSS is 0.69.

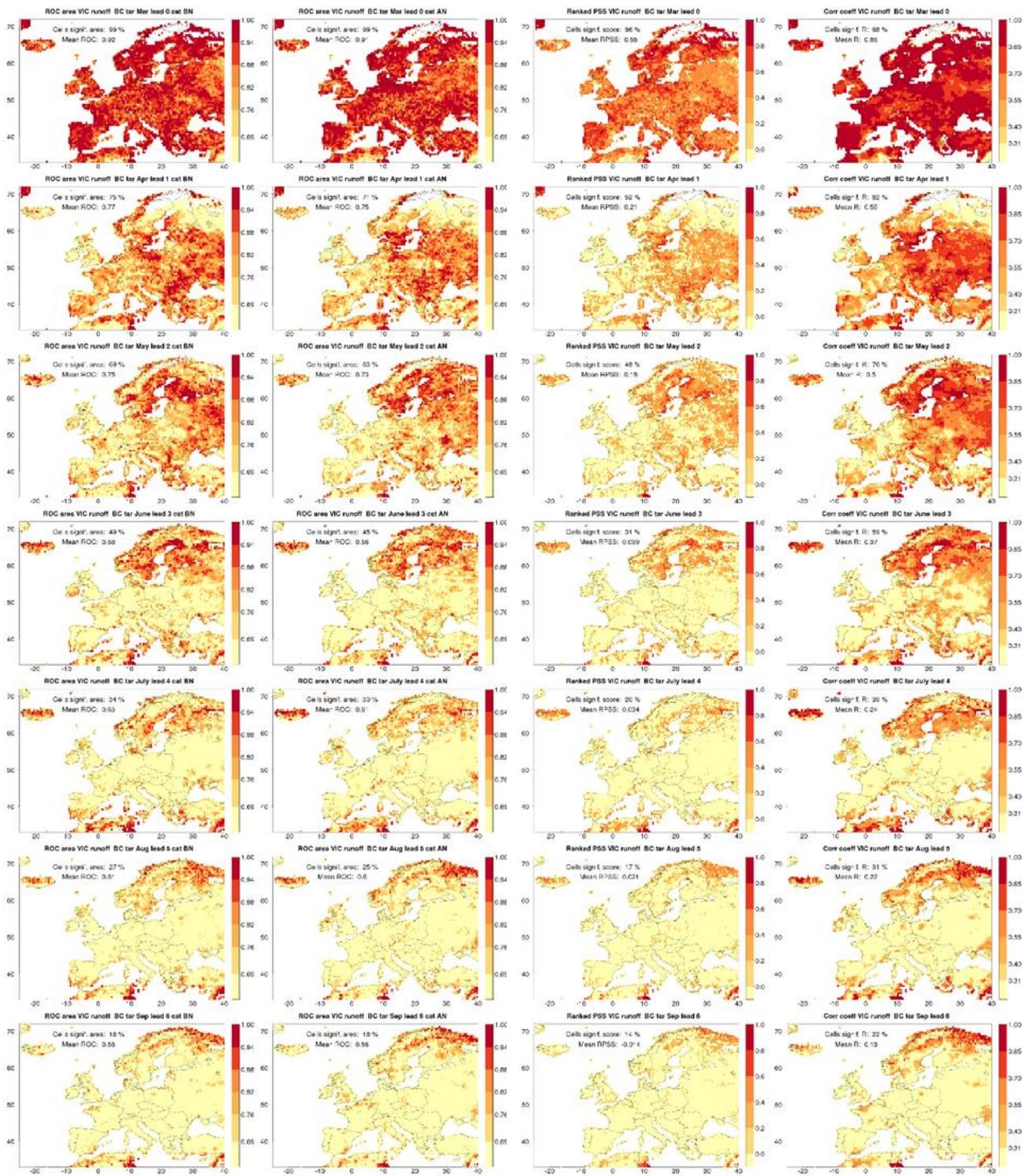
# Initialisation January



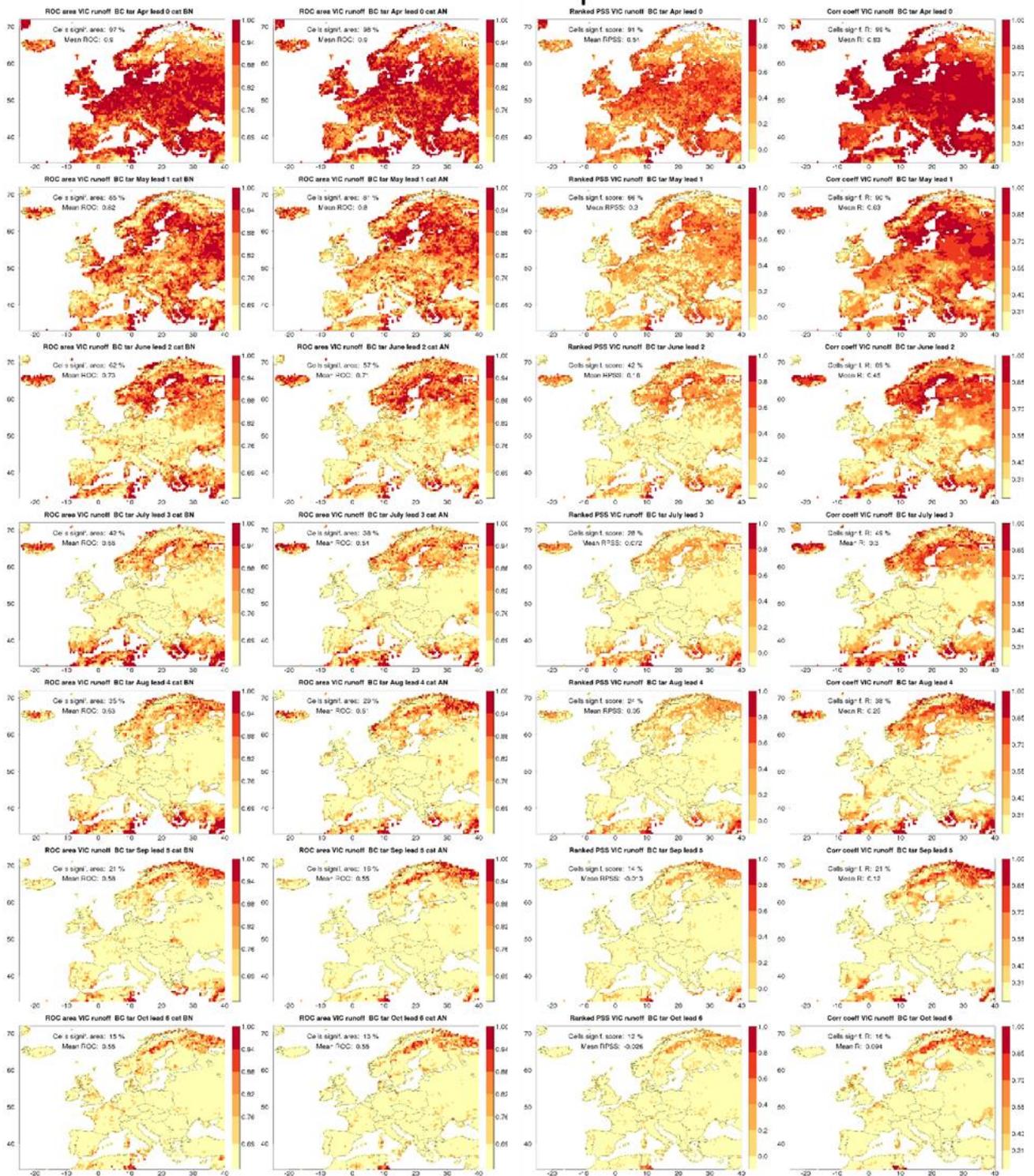
# Initialisation February



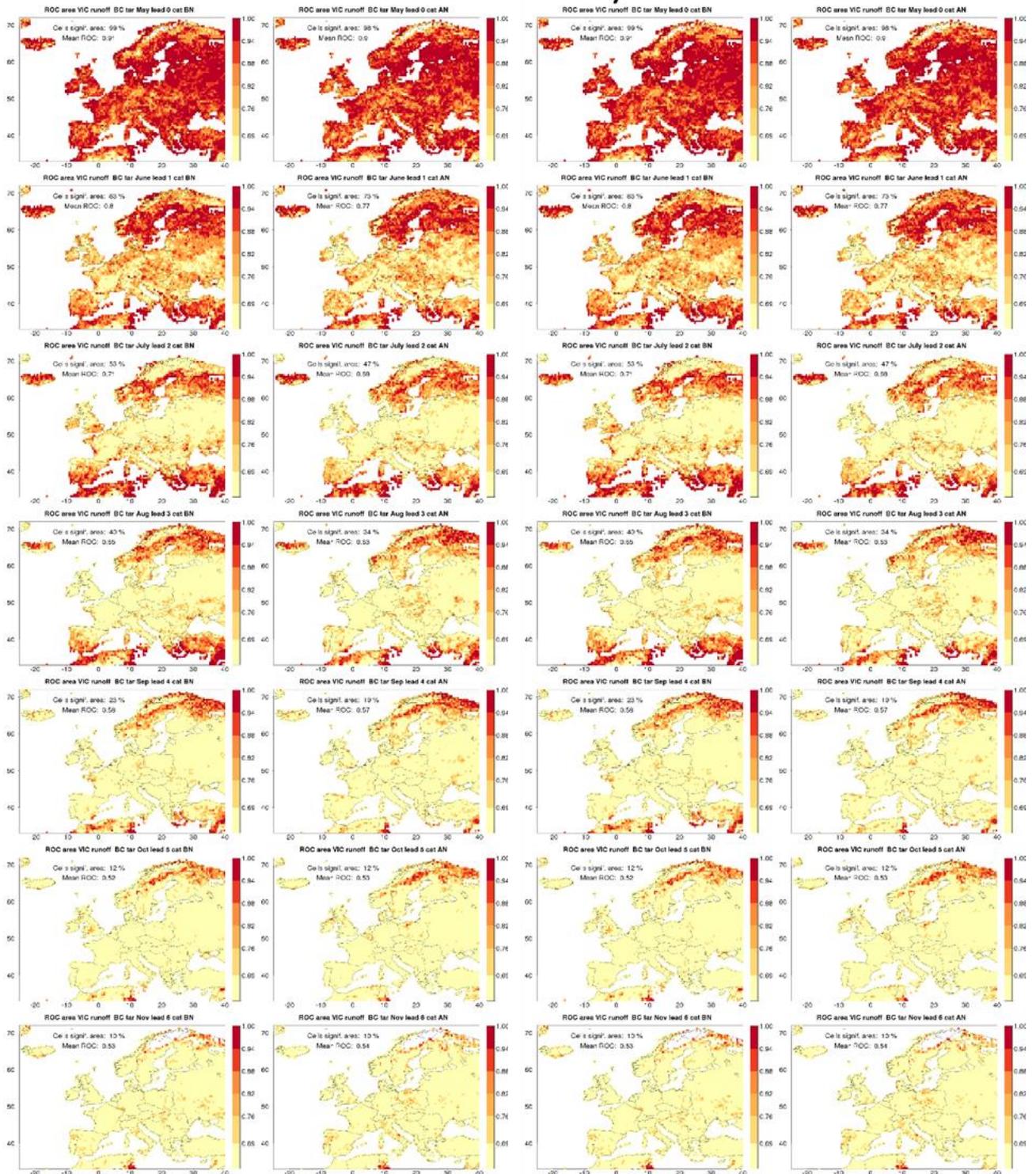
# Initialisation March



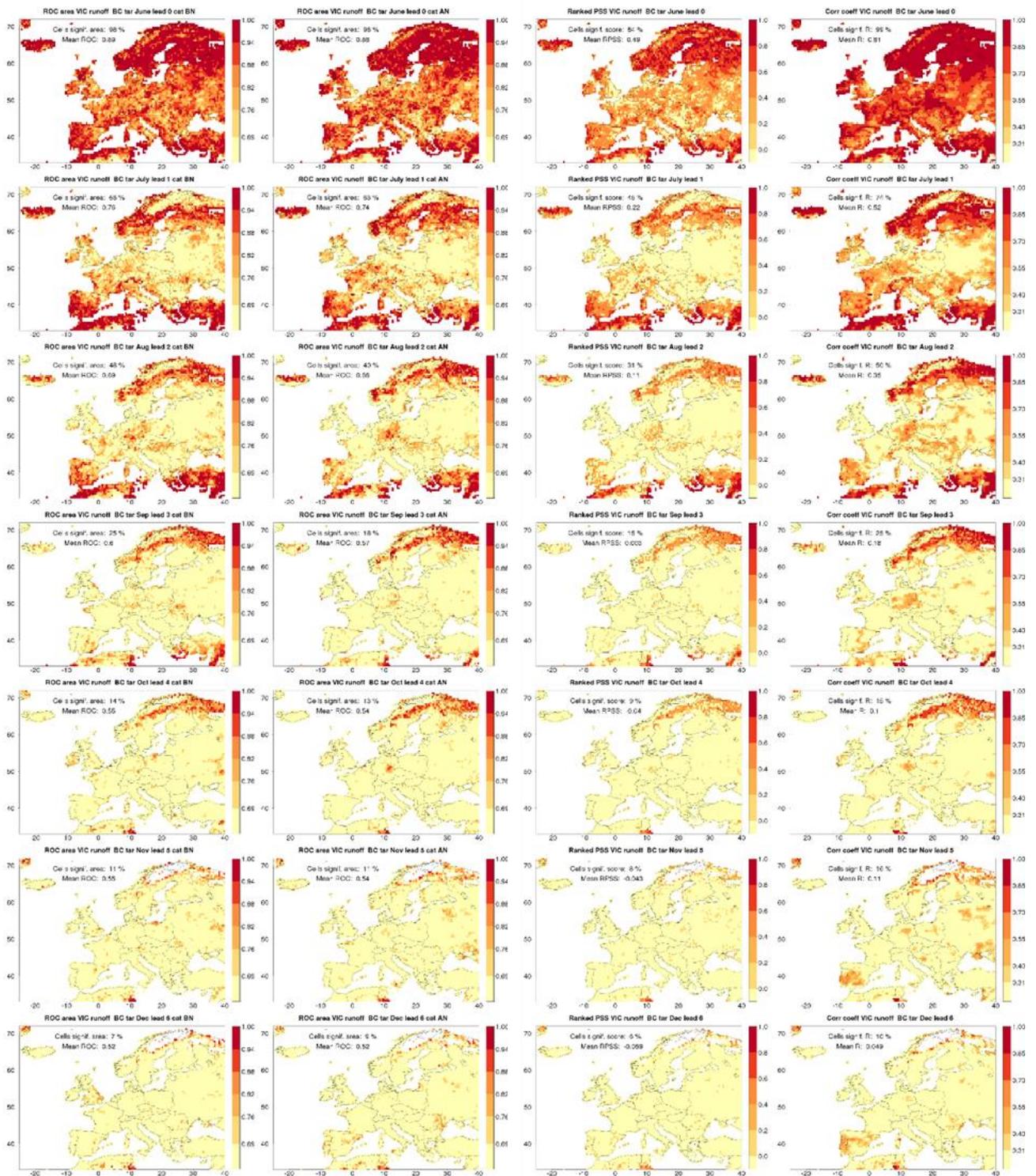
# Initialisation April



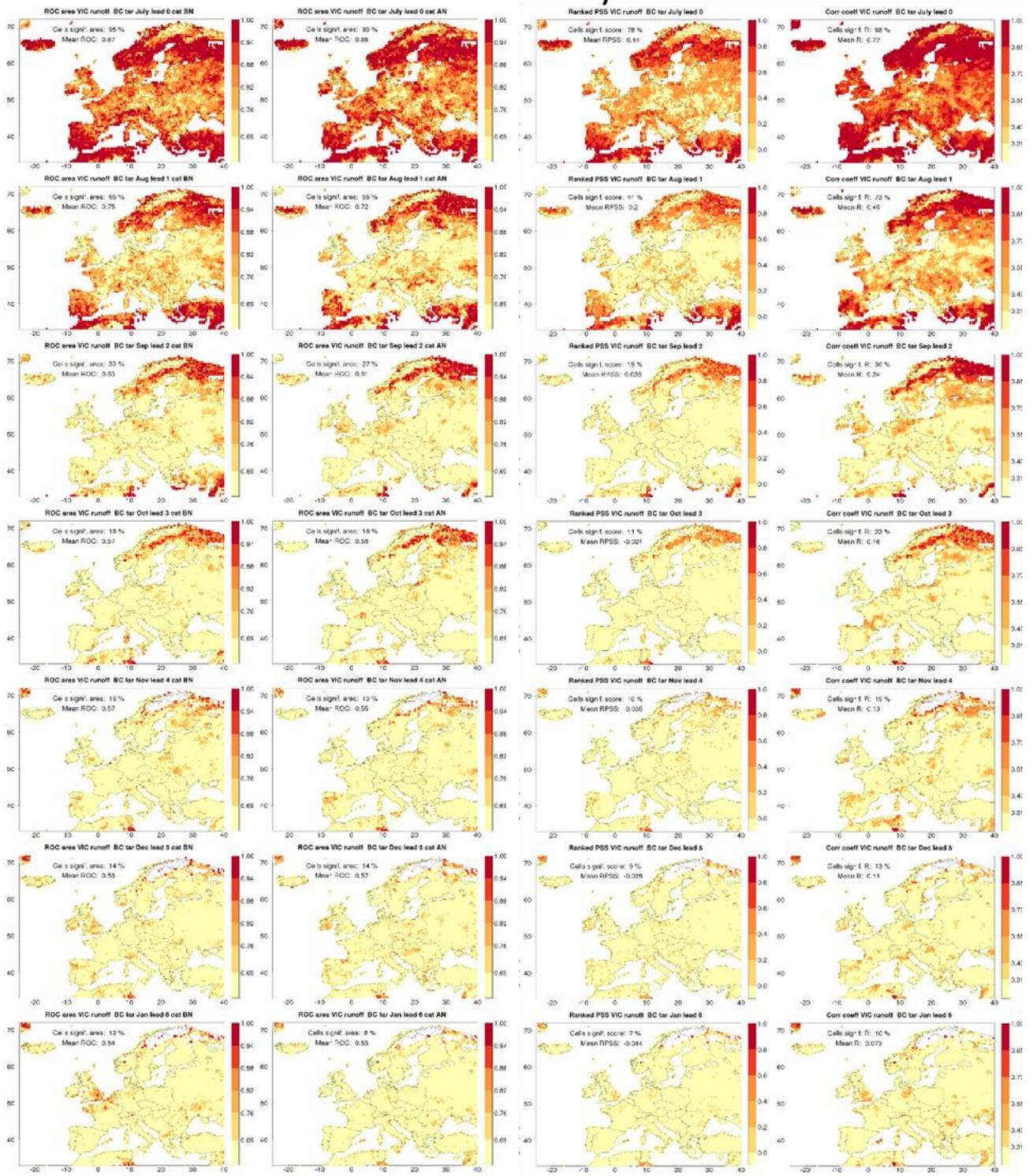
# Initialisation May



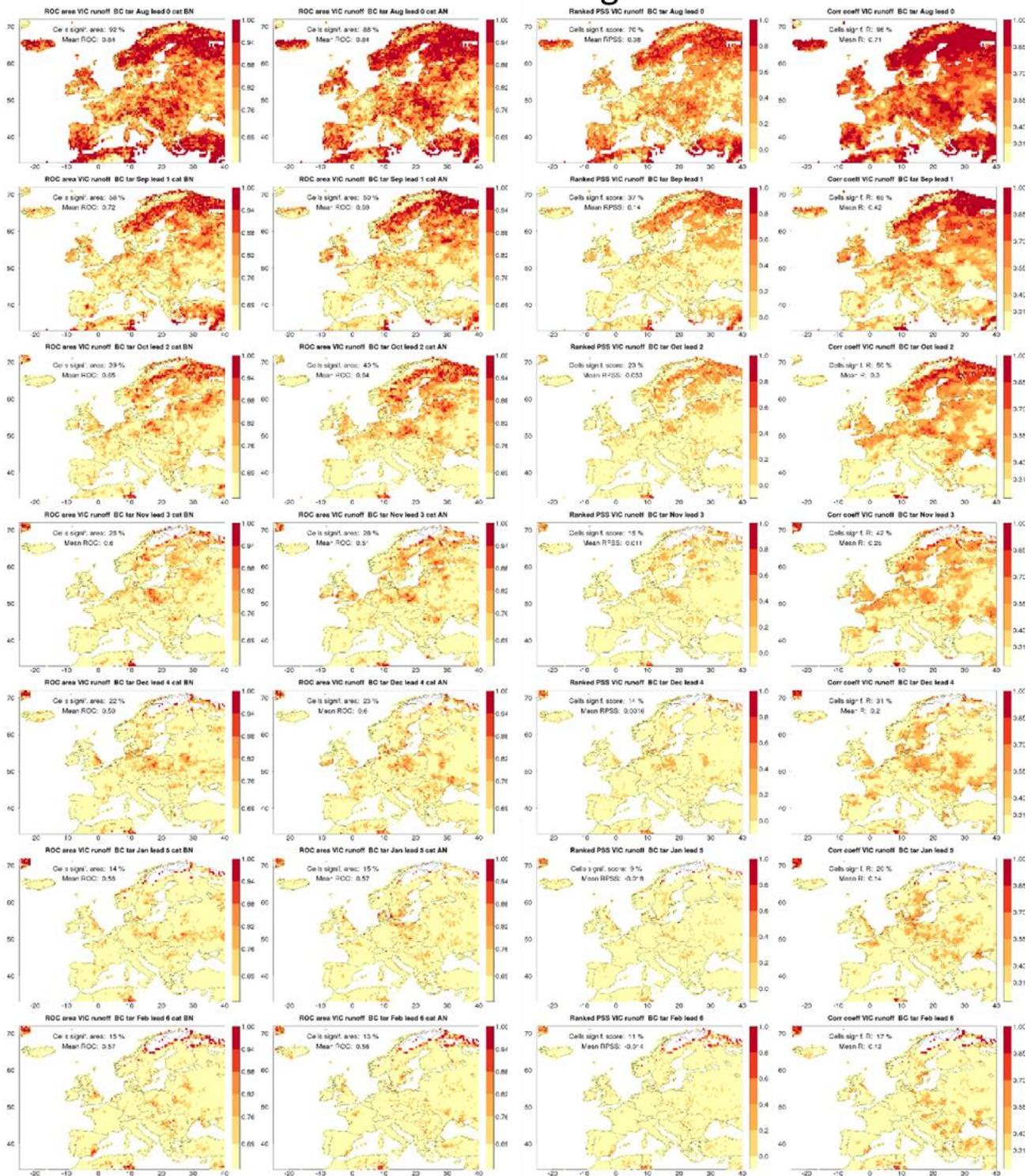
# Initialisation June



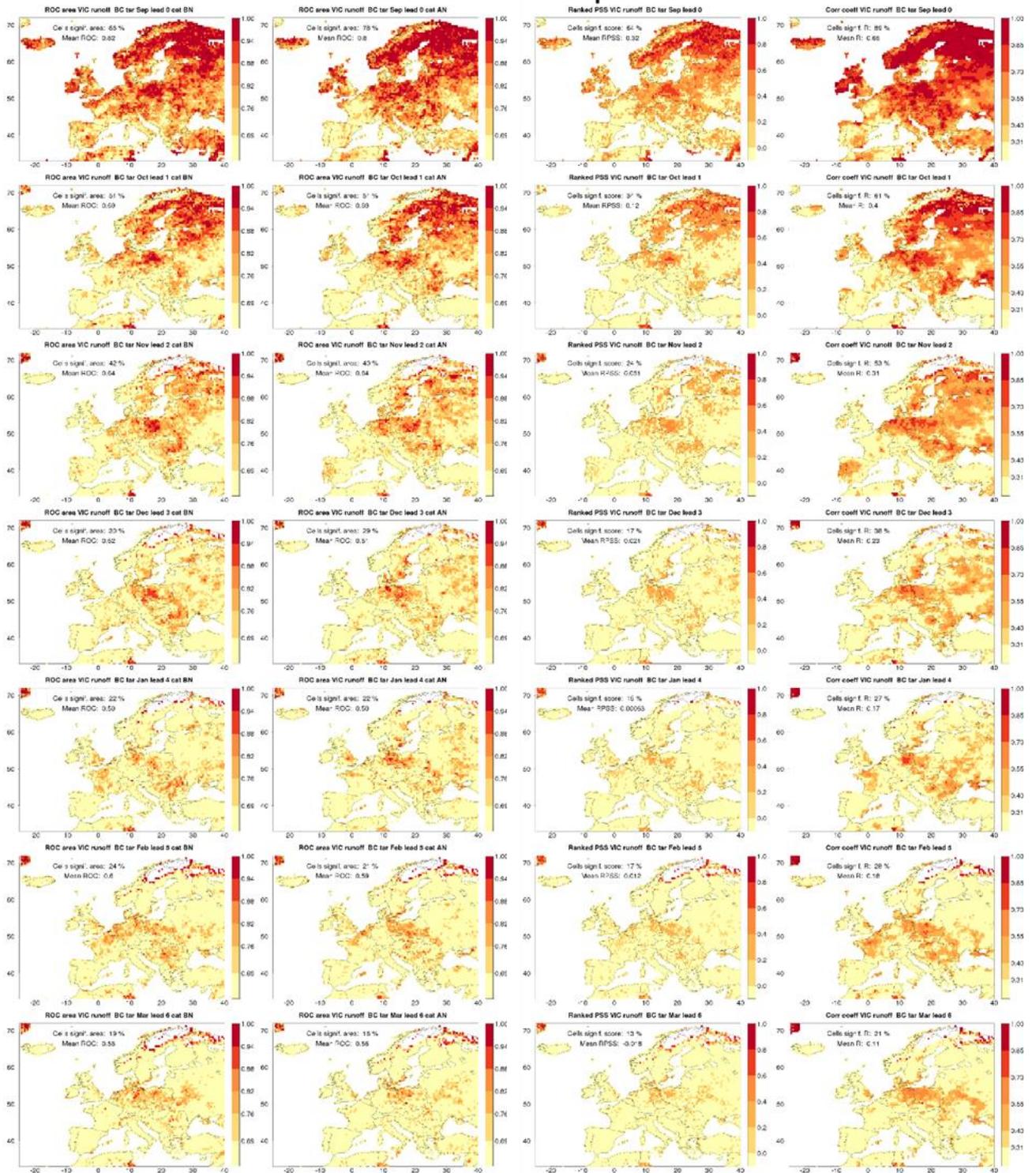
# Initialisation July



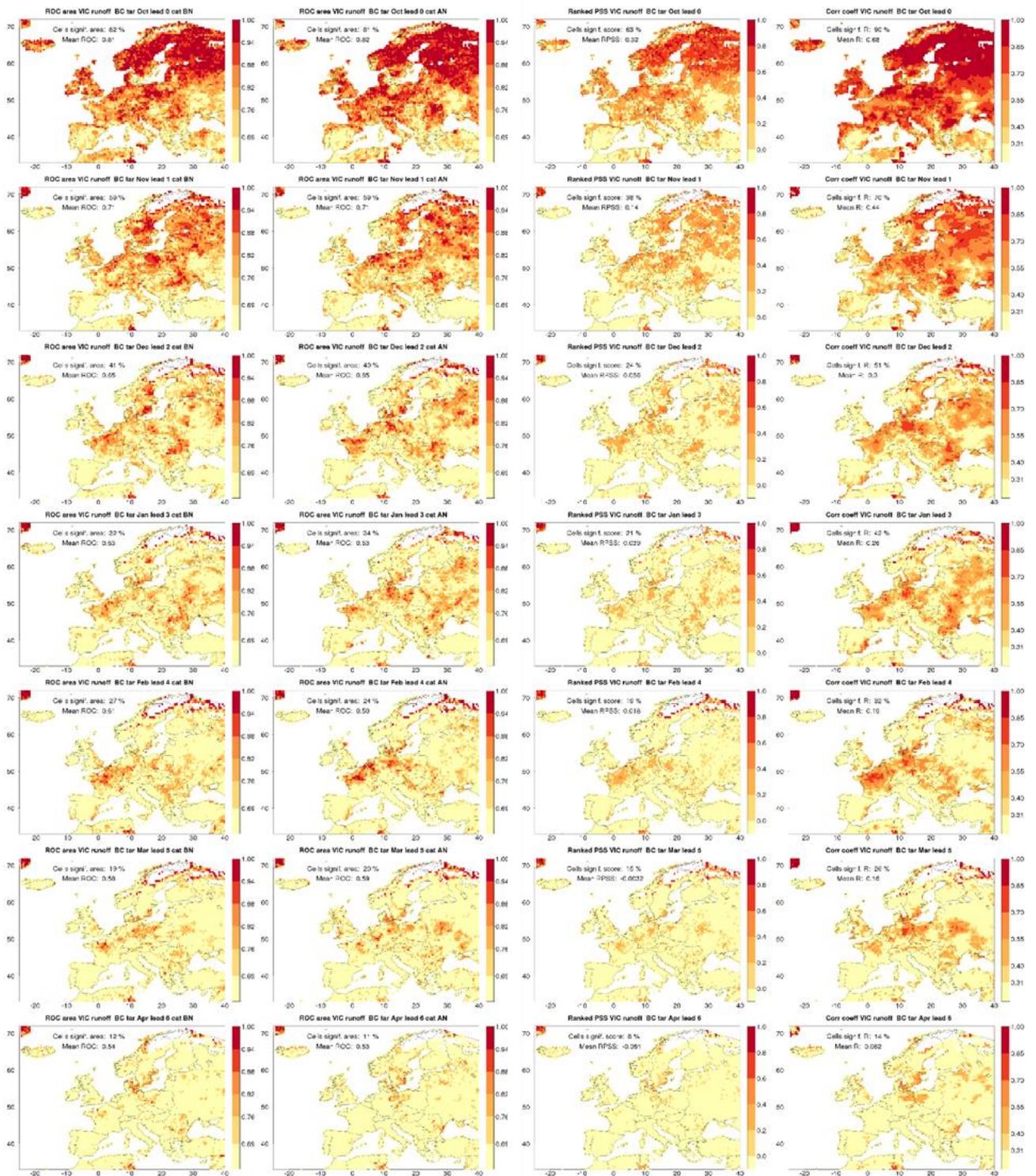
# Initialisation August



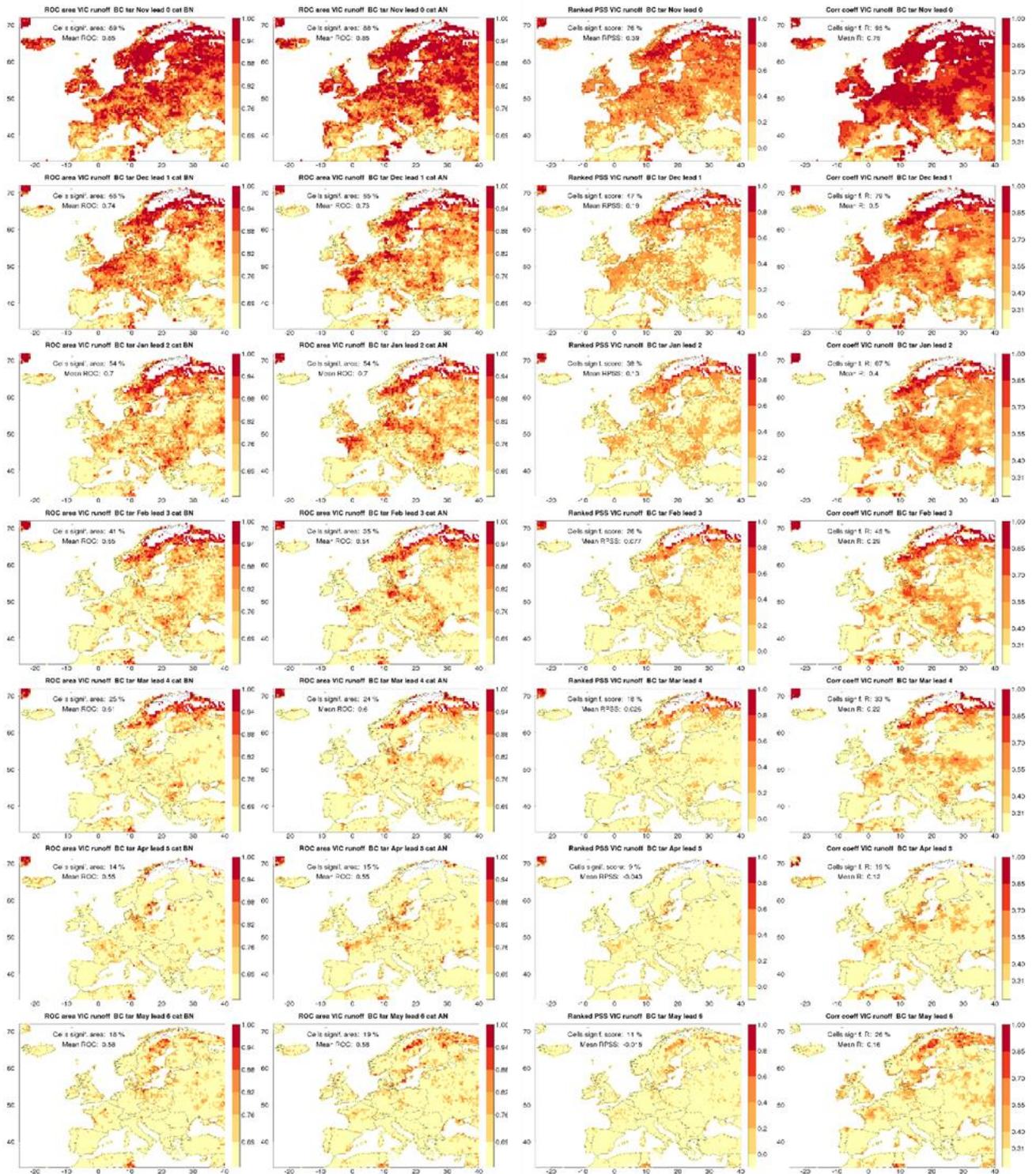
# Initialisation September



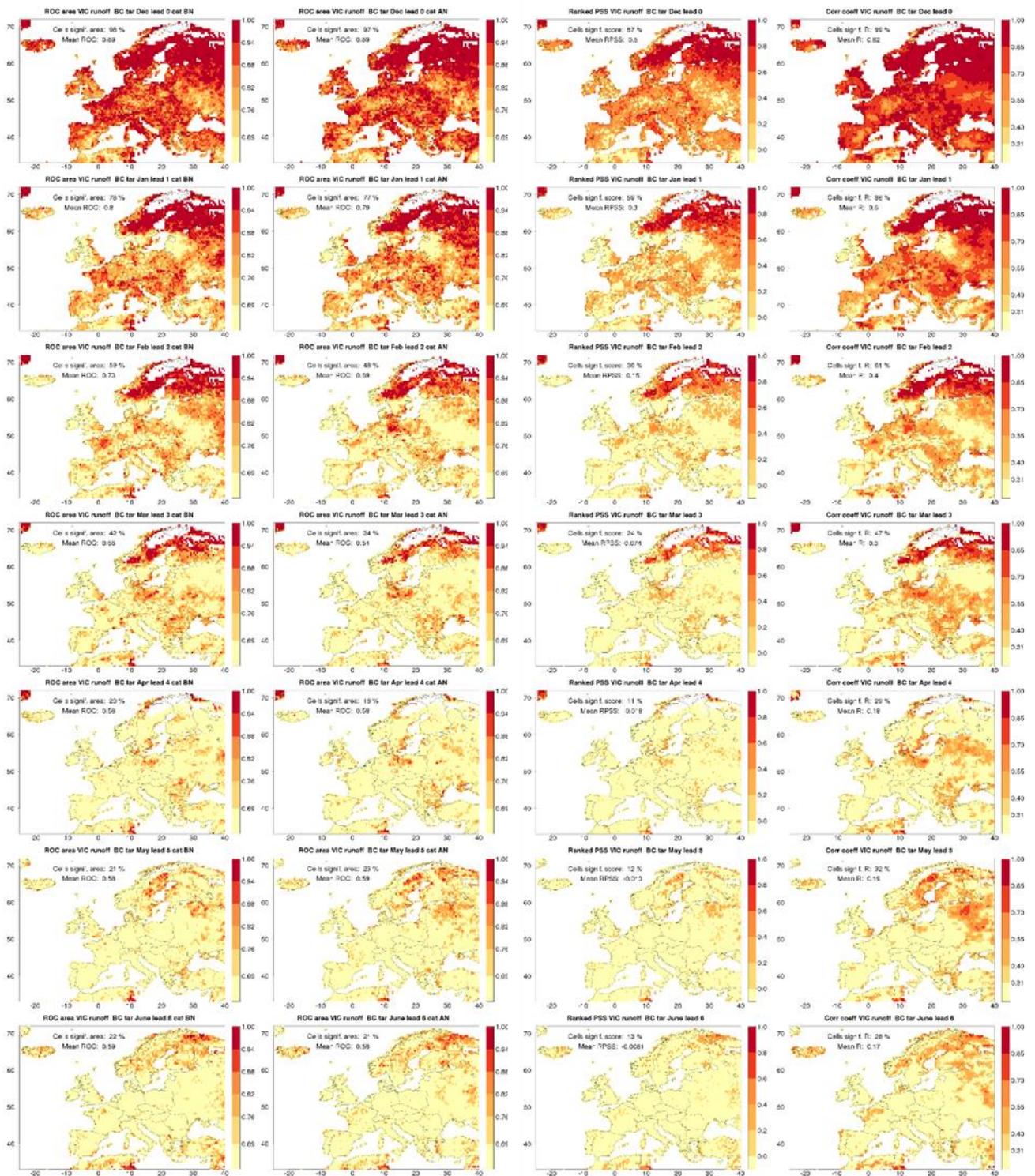
# Initialisation October



# Initialisation November



# Initialisation December



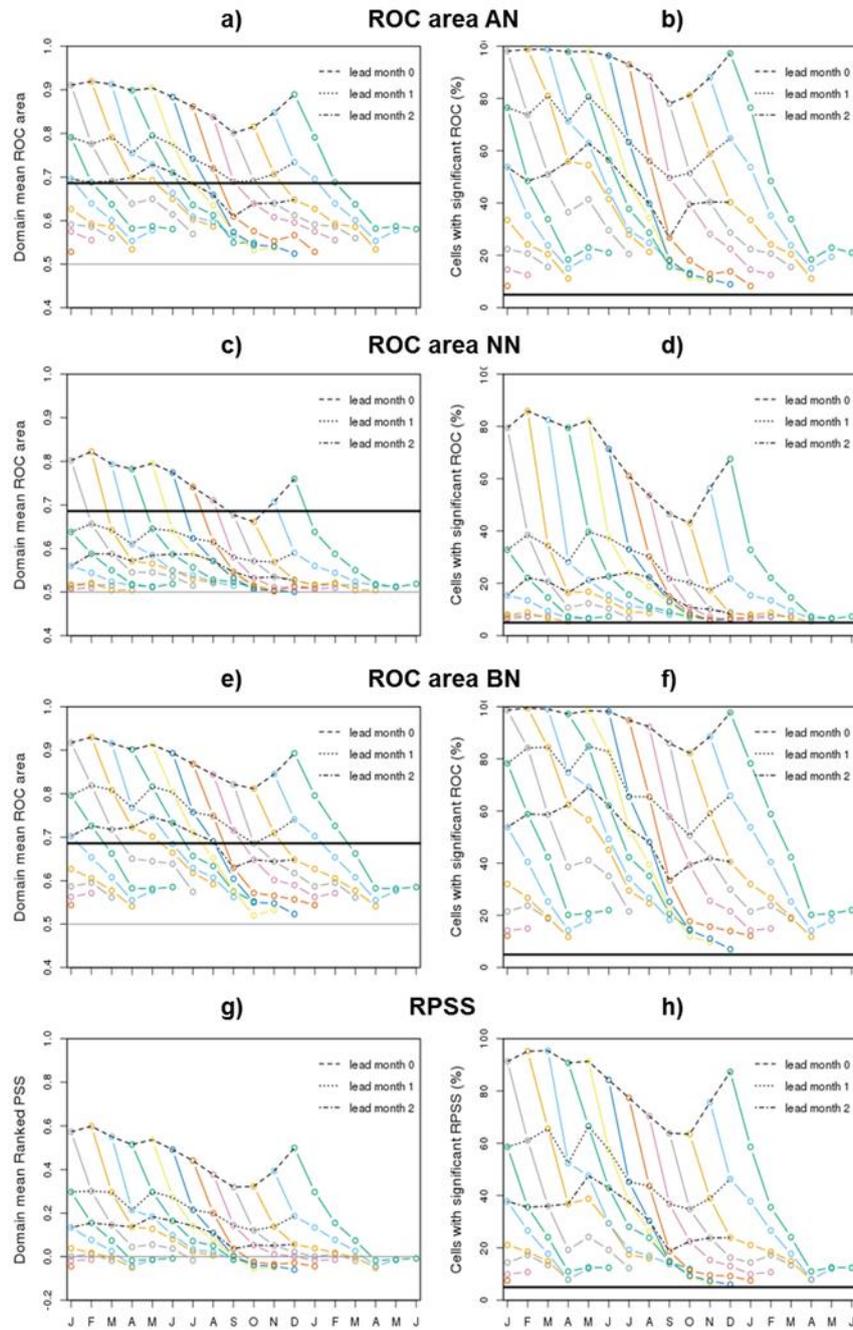


Figure S2: This figure supplements Fig. 5 (which for convenience is reproduced here at the top). It shows four different metrics (from top to bottom: R, ROC area for AN and BN tercile, and RPSS) and two statistics of those metrics (domain mean value of the particular metric at left, and percentage of cells with significant skill in that metric at right) as a function of initialisation month and lead time. The thick horizontal lines in the left-hand column give the threshold between significant and insignificant skill for *individual* cells (0.31 for R and 0.69 for ROC area). The thick horizontal lines in the right-hand column give the expected fraction of cells with significant skill, in the case that the hindcasts have no skill at all (5% for all metrics). More explanation is given in the caption of Fig. 5.

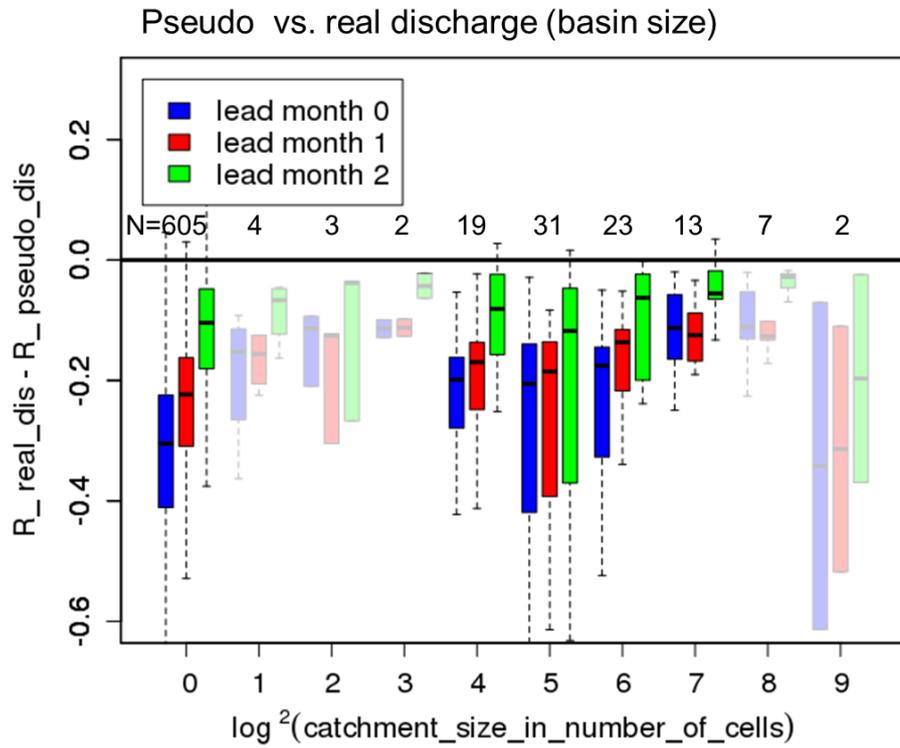


Figure S3 Similar to figure 6d, this figure plots the difference between pseudo and real skill (using correlation coefficient as metric) as a function of basin size. The left most box group is for the small, i.e. subgrid basin stations, the others are for the larger basins. The number of stations in each bin are given above the 0 line. Bins with fewer than 10 stations have been masked. Each box with whiskers represents minimum, lower quartile, median, upper quartile and maximum, respectively.

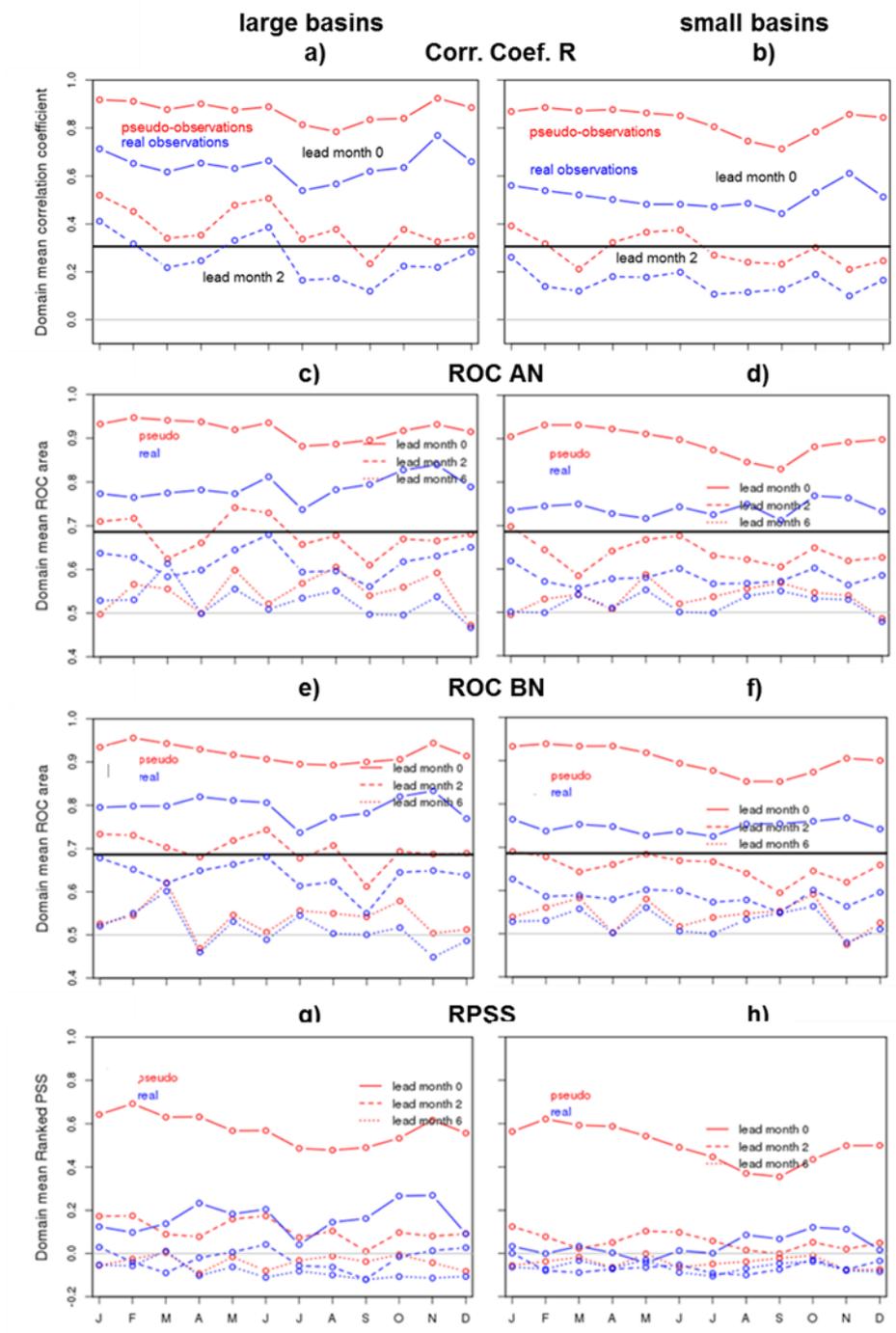


Figure S4: This figure supplements Fig. 8 (copied here in a ,b) for convenience). It compares verification of discharge with pseudo- (red) and real (blue) observations, in terms of the annual cycle of different metrics for large basins (left-hand column) and small basins (right-hand column). The metrics are, from top to bottom, domain mean ROC area for the AN (a and d) and the BN tercile (b and e), domain-mean RPSS (c and f) and the percentage of cells with significant skill in terms of R (g and k), the ROC area for the AN (h and l) and the BN tercile (i and m), and the RPSS (j and n). All panels depict annual cycles for lead months 0, 2 and 6. The thick horizontal lines give the threshold between significant and insignificant skill for individual

cells in (a, b, d and e) and the expected fraction of cells with significant skill, in the case that the hindcasts have no skill at all (g-n).

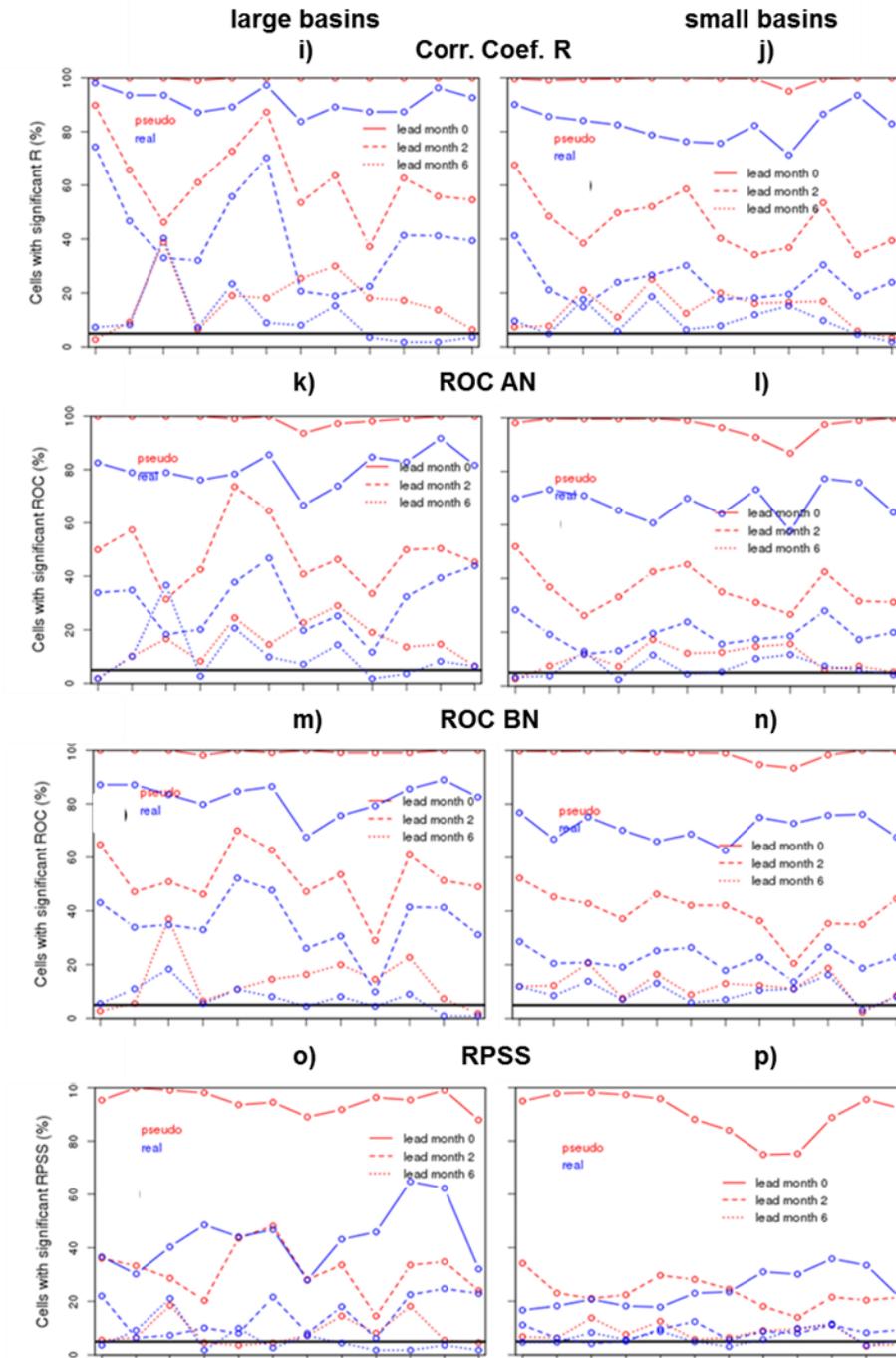


Figure S4 continued